

Université de Lausanne - Faculté de Médecine

Institut de Physiologie

Directeur : Professeur A. Fleisch

Lauf Medical Library

Le calcul des moyennes en biologie

Thèse

présentée à la Faculté de Médecine de l'Université de Lausanne
pour l'obtention du grade de docteur en médecine

par

Jean-François Cuendet

médecin diplômé de la Confédération suisse



Digitized by the Internet Archive
in 2026

Université de Lausanne - Faculté de Médecine

Institut de Physiologie

Directeur : Professeur A. Fleisch

Le calcul des moyennes en biologie

Thèse

présentée à la Faculté de Médecine de l'Université de Lausanne
pour l'obtention du grade de docteur en médecine

par

Jean-François Cuendet

médecin diplômé de la Confédération suisse

UNIVERSITÉ DE LAUSANNE

Faculté de Médecine

Le Conseil de la faculté de médecine de l'Université de Lausanne, après avoir pris connaissance des rapports favorables de MM. les professeurs Fleisch et Demole et sans se prononcer sur les opinions du candidat, autorise l'impression de la thèse intitulée :

Le calcul des moyennes en biologie
présentée par Monsieur Jean-François Cuendet pour l'obtention du grade de Docteur en médecine.

Lausanne, le 24 septembre 1947.

Le Doyen :

Prof. Dr DECKER

A mon père

*Nous tenons à exprimer à Monsieur le
Professeur A. Fleisch notre très vive recon-
naissance pour la grande bienveillance avec
laquelle il a dirigé ce travail.*

Introduction

Quelle est la raison de l'importance du calcul statistique en biologie ?

Les mesures biologiques d'un même phénomène effectuées dans des conditions identiques ne sont pas constantes. Elles changent d'un sujet d'expérience à l'autre. Elles peuvent varier également pour le même sujet au cours d'expériences successives. Alors que pour les sciences exactes une seule mesure suffit souvent pour caractériser un phénomène, cette *variabilité* nécessite une multiplicité de mesures en biologie. Comment donner une expression quantitative à l'ensemble de ces données variables, comment les caractériser, comment comparer entre eux deux collectifs de telles données ? Le calcul statistique permet de résoudre ces problèmes, aussi comprenons-nous pourquoi il devrait occuper une place importante en biologie.

En présence d'un collectif de mesures variables, la statistique nous permet de calculer les caractéristiques suivantes :

- la moyenne
- la dispersion

et pour ceux qui aiment la précision mathématique

- l'indice d'asymétrie et celui de probabilité des extrêmes.

La moyenne est la caractéristique la plus simple d'un collectif de mesures variables.

Deux formes de moyennes sont d'usage courant en statistique : la médiane et la moyenne arithmétique.

Si l'on classe les données par ordre de valeur croissante, la *médiane* est la valeur centrale, celle qui se trouve au milieu de toutes les autres. Moyenne approximative, la médiane a l'avantage de réduire les calculs au minimum.

C'est la *moyenne arithmétique* qui est presque toujours appliquée en biologie. Diviser la somme des valeurs par leur nombre, voilà une opération précise et d'un calcul facile.

La *moyenne géométrique* consiste à effectuer le produit des valeurs entre elles puis à en extraire la racine d'indice égal à leur nombre. Elle est utilisée seulement dans des cas exceptionnels, qui sont presque tous étrangers à la biologie.

Yule et Kendall (21) prétendent qu'il faudrait appliquer la moyenne géométrique chaque fois que les variations d'une quantité sont proportionnelles à cette quantité elle-même. Ainsi le nombre d'habitants d'un pays. Supposons qu'on connaisse le chiffre d'une population à deux époques déterminées. Pour calculer le chiffre de population d'une époque intermédiaire à ces deux époques, il y a lieu d'appliquer la moyenne géométrique.

De Moncetz (17) montre que pour calculer le taux d'intérêt moyen R d'une somme A placée pendant plusieurs années à des taux variables $r_1, r_2 \dots$ etc., il faut prendre la moyenne géométrique des expressions $(1+r_1), (1+r_2) \dots$ etc., soit :

$$1+R = \sqrt[n]{(1+r_1) \cdot (1+r_2) \dots (1+r_n)}$$

La géométrie ordinaire fournit un autre exemple de moyenne géométrique cité par le même auteur : si un parallépipède rectangle a pour hauteur a , pour longueur b et pour largeur c , le cube du même volume a pour arête la racine cubique du produit $a \cdot b \cdot c$. Autrement dit, l'arête du cube est égale à la moyenne géométrique des trois grandeurs a, b et c , et non à leur moyenne arithmétique.

La moyenne géométrique est couramment utilisée pour

les calculs de nombres-indices. *Le British Board of Trade* l'applique aux indices des prix de gros. *Donzallaz* (5) au calcul des prix et des degrés d'occupation, un rapport entre l'offre et la demande de travail.

En biologie, dans une étude sur la croissance des cultures d'algues, *Bristol Roach* (3) applique implicitement la moyenne géométrique puisqu'il effectue ses calculs sur les logarithmes décimaux du volume occupé par les cultures. Mais ce n'est qu'avec les travaux de *Fleisch* (9), *Fleisch et Tripod* (8), *Fleisch et Posternak* (10) que l'avantage de la moyenne géométrique sur la moyenne arithmétique a été reconnu dans de nombreux domaines de la biologie.

Les deux indications majeures d'application de la moyenne géométrique sont la présence d'un collectif constitué par des rapports et celle de données dépendant de la sensibilité du sujet d'expérience.

Pourquoi appliquer la moyenne géométrique aux rapports ? Pour des raisons de simple logique : il est nécessaire que la moyenne de rapports calculés dans un sens soit l'inverse de la moyenne de ces mêmes rapports calculés dans l'autre sens.

Reprenons l'exemple de *Fleisch et Tripod* (8). Comparant l'action biologique de deux substances A et B, on mesure la dose nécessaire de chaque substance pour provoquer le même effet. Pour chaque essai on effectue le rapport entre la dose de l'une des substances et celle de l'autre substance considérée comme standard avec la valeur arbitraire de 1. Les doses suivantes ont provoqué le même effet biologique :

	<u>$A : B = x : 1$</u>	<u>$B : A = y : 1$</u>
8 mg A = 2 mg B	4 : 1	0,25 : 1
2 mg A = 1 mg B	2 : 1	0,5 : 1
2 mg A = 2 mg B	1 : 1	1 : 1
4 mg A = 2 mg B	<u>2 : 1</u>	<u>0,5 : 1</u>
Moyenne arithmétique	2,25 : 1	0,562 : 1

La moyenne arithmétique des rapports $A : B = 2,25$ n'est pas l'inverse de celle des rapports $B : A = 0,562$. Voilà qui est illogique ! Par contre la moyenne géométrique, qui est

pour le rapport A : B de 2 mg et pour le rapport B : A de 0,5 mg, n'a pas cet inconvénient.

De Moncetz (17) fait la même remarque à propos des nombres-indices, qui sont des rapports entre un chiffre actuel et celui d'une année antérieure choisie comme base. L'indice des prix de gros, l'indice du coût de la vie, les indices boursiers etc. sont des moyennes d'indices, pondérés ou non. L'indice calculé dans un sens devrait être l'inverse de celui calculé dans l'autre sens. Autrement dit, le produit de deux indices à bases inversées devrait toujours être égal à l'unité. Or les indices calculés par la moyenne arithmétique ne répondent pas à cette condition. A l'appui de ces considérations théoriques sur la moyenne de rapports, nous exposerons plus loin un exemple pratique.

La moyenne géométrique paraît également indiquée pour les collectifs de données dépendant de la *sensibilité* du sujet d'expérience. Ainsi, *Fleisch et Posternak* (10) ont appliqué la moyenne géométrique à la courbe d'adaptation de la vision à l'obscurité. Nous examinerons plus loin des statistiques de sensibilité de la peau chez l'homme, de sensibilité de la pulpe dentaire du lapin, de dose mortelle de strophantine pour la grenouille. Pour ces 3 statistiques, la concordance avec la courbe de Gauss idéale a été le critère de l'avantage de l'une ou l'autre moyenne. Dans ces 3 cas la moyenne géométrique s'est révélée préférable.

Parallèlement à l'étude concrète de ces statistiques, des considérations théoriques nous amènent à préférer la moyenne géométrique dans les collectifs mesurant la sensibilité d'un sujet d'expérience. *Fleisch et Tripod* (8) ont montré que la dispersion de ces données ne se fait pas symétriquement en plus ou en moins autour d'une valeur moyenne. La dispersion se fait dans le sens d'un multiple ou d'une fraction de la valeur moyenne. Par exemple la dose mortelle d'un médicament peut être habituellement de 1 mg. Elle pourra être de 2, voir de 3 mg pour des sujets très résistants. Les individus très sensibles correspondants mourront avec $\frac{1}{2}$ ou $\frac{1}{3}$ mg, alors que si la dispersion se faisait symétriquement en + ou en —, une dose

mortelle de 2 mg devrait avoir comme pendant une dose nulle, et une dose de 3 mg aurait comme pendant une dose négative, ce qui est un non-sens! Les mêmes considérations sont justifiées pour un phénomène sensoriel quelconque.

Mais voici encore d'autres arguments en faveur de la moyenne géométrique. Les phénomènes sensibles obéissent à la loi de *Weber*, disant que l'accroissement perceptible minimum d'excitation ΔE est proportionnel à l'excitation préexistante E

$$\frac{\Delta E}{E} = \text{constante}$$

Par intégration et en supposant que la différence de sensation ΔS est la même pour tous les degrés d'excitation préexistante, *Fechner* a établi sa célèbre loi psychophysique :

$$S = C. \log. \text{ nat. } E$$

où S désigne la sensation, E l'excitation extérieure et où C est une constante. Dans tout phénomène sensoriel, il importe de distinguer un phénomène physique, l'excitation et un phénomène purement biologique, la sensation. Comme il est impossible de mesurer l'intensité de la sensation, nous en sommes réduits à mesurer l'intensité de l'excitation. Si la loi de *Fechner* a une réalité, pour le moins approximative comme l'ont prouvé de nombreux phénomènes, nous pouvons faire les réflexions suivantes. Ce qui importe pour la discrimination entre deux sensations, c'est la différence d'intensité des sensations. Nous pouvons imaginer que dans le monde sensoriel psychique il existe une échelle arithmétique à laquelle répondent dans le monde physique de l'excitation les logarithmes des intensités de l'excitation. C'est ainsi qu'il est indiqué de ne pas calculer les moyennes arithmétiques des intensités d'excitation, mais les moyennes arithmétiques de leurs logarithmes c. à d. la moyenne géométrique. Car prendre la moyenne arithmétique des logarithmes des mesures équivaut à prendre la moyenne géométrique des mesures elles-mêmes.

Les modes de variation de l'excitation et de la sensation

se déduisent mathématiquement par la loi de Fechner. Si la sensation S varie suivant le mode arithmétique ($S \pm a$) c. à d. symétriquement en $+$ ou en $-$, l'excitation E doit varier suivant le mode géométrique. E multiplié, respectivement divisé par b , $E \times b$; $E \div b$. L'expression mathématique en est $E \cdot e^{\pm b}$

Transformons la loi de Fechner $S = C \text{ Log. } E$. Ajoutons à S un coefficient arithmétique de variation $\pm a$. Ajoutons cette même quantité également dans le deuxième membre de l'équation pour rétablir l'égalité. L'équation prend la forme suivante

$$S \pm a = C \text{ Log } E \pm a$$

$$\text{or } \pm a = \text{Log } e^{\pm a} = C \text{ Log } e^{\pm \frac{a}{C}}$$

remplaçons $\pm a$ par cette nouvelle valeur dans l'équation de Fechner

$$S \pm a = C \text{ Log } E + C \text{ Log } e^{\pm \frac{a}{C}}$$

d'où

$$S \pm a = C \text{ Log } \left[E \cdot e^{\pm \frac{a}{C}} \right]$$

ce qu'on peut exprimer également

$$S \pm a = C \text{ Log } \left[E^{\times \frac{a}{C}} \right]$$

L'introduction d'un coefficient arithmétique $\pm a$ à l'intérieur de la parenthèse logarithmique détermine les variations de l'excitation dans le sens d'un multiple ou d'une fraction d'elle-même. Les variations correspondantes de la sensation se font selon le mode arithmétique $S \pm a$.

En résumé, l'analyse de statistiques concrètes, l'étude des variations de l'excitation, les lois de Weber et de Fechner, nous incitent à appliquer la moyenne géométrique aux mesures physiques où la sensibilité individuelle joue un rôle.

Il y a certainement encore beaucoup de phénomènes où la moyenne arithmétique n'est pas indiquée.

Enfin certains collectifs comme celui de *Fleisch et Tripod* (8) sur le temps de saignement du lapin, ne relèvent ni de l'une ni de l'autre moyenne. Seule une formule intermédiaire, telle que l'ont décrite ces auteurs, semble satisfaisante.

B. Etude de 11 collectifs biologiques

Pour comparer pratiquement les avantages de la moyenne arithmétique et ceux de la moyenne géométrique, nous avons appliqué à 11 collectifs biologiques la statistique arithmétique et la statistique géométrique. Nous n'avons pas seulement comparé les moyennes, mais aussi la dispersion et les autres caractéristiques de la courbe. Cette comparaison se fait graphiquement avec la courbe de Gauss idéale. Les exemples ont été choisis dans les domaines les plus divers de la biologie : index de nutrition des enfants, sensibilité de la pulpe dentaire du lapin, seuil de sensibilité cutanée, dose mortelle de strophantine pour la grenouille, durée d'une course à pied, taux de cholestérol dans le sang, taille des conscrits, périmètre du thorax, distance d'un lancer, réfraction cristallinienne, formule d'Arneth.

Le nombre de cas observées a été choisi aussi grand que possible pour avoir un matériel homogène. Il varie pour nos statistiques de 149 à 31.806. Nous avons recherché pour chaque collectif les mesures originales, pas encore élaborées par la statistique comme elles le sont presque toujours dans une publication.

1. METHODE

Nous avons adopté pour nos statistiques les méthodes de *Fueter* (11). Pour chaque statistique, nous avons effectué les calculs d'une part sur les données originales, d'autre

part sur leurs logarithmes. Dans ce dernier cas, prendre la moyenne arithmétique des logarithmes équivaut à prendre la moyenne géométrique des nombres. Qu'il s'agisse des nombres originaux ou de leurs logarithmes, les méthodes de calcul sont les mêmes.

Le matériel est réparti en classes équidistantes dont le nombre est choisi arbitrairement en tenant compte de la dispersion (σ) et du nombre des cas observés (n). On appelle w l'étendue d'une classe. Pour chaque classe on compte le nombre de cas observés $f(x)$. Puis on numérote les classes, en attribuant le numéro $x=0$ à la classe qui comprend le plus grand nombre de cas. Les autres classes sont numérotées à partir de celle-là : les classes comprenant des valeurs inférieures reçoivent successivement les numéros $x=-1, x=-2, x=-3$ etc., les classes supérieures, $x=1, +2, +3$ etc. On multiplie ensuite le nombre des cas de chaque classe $f(x)$ avec la première puissance du numéro d'ordre de la classe (x). La somme de ces produits $\sum x.f(x)$ nous permet d'établir la moyenne arithmétique des données de la statistique. Le milieu de la classe $x=0$, désigné M_0 , nous donne déjà une première approximation de la moyenne arithmétique. La moyenne arithmétique exacte M des résultats se calcule à l'aide des formules suivantes :

$$M = M_0 + b \quad b = \frac{w}{n} \sum x f(x) \quad \text{où } b \text{ est un facteur}$$

de correction.

La somme des produits entre le carré du numéro d'ordre d'une classe et le nombre de cas de cette classe $\sum x^2.f(x)$ nous permet de calculer la dispersion (σ) à l'aide de la formule suivante :

$$\sigma = \sqrt{\frac{w^2}{n} \sum x^2.f(x) - b^2}$$

où b est le facteur de correction que nous venons de voir plus haut. L'erreur moyenne (ε) de la moyenne arithmétique se calcule à l'aide de la formule suivante :

$$\varepsilon = \frac{\sigma}{\sqrt{n}}$$

La représentation graphique de la statistique est très précieuse. Elle permet d'apprécier visuellement la concordance entre la courbe idéale de Gauss et celle des résultats expérimentaux. Elle l'est encore davantage si l'on compare deux courbes, comme nous l'avons fait pour les courbes des statistiques arithmétique et géométrique avec la courbe idéale de Gauss. La meilleure concordance de l'une ou de l'autre apparaît au premier coup d'œil.

Autrefois, on considérait une asymétrie de la courbe de Gauss comme impossible. On était d'avis qu'une pareille statistique renfermait une erreur systématique. Plus tard on a reconnu que la courbe symétrique de Gauss n'est qu'un cas spécial, idéal, et on a développé la théorie mathématique des courbes asymétriques. *Charlier* (4) insiste sur le fait qu'une asymétrie de la courbe de Gauss prouve que la moyenne arithmétique ne représente pas la moyenne adéquate. Des courbes asymétriques peuvent souvent être transformées en courbes de Gauss symétriques en employant pour l'axe des x , non pas la progression arithmétique, mais une fonction de x comme par exemple $\sqrt{x}$ ou $\log x$. Nous pouvons donc dire qu'en employant pour l'axe des x une fonction arithmétique et en obtenant ainsi une courbe asymétrique, la moyenne arithmétique n'est pas adéquate. Par contre si nous employons pour l'axe des x la progression logarithmique et si nous obtenons ainsi une courbe de Gauss symétrique, cela nous prouve l'avantage de la moyenne géométrique.

Comme il serait difficile de construire une nouvelle courbe de Gauss pour chaque dispersion (σ), *Fueter* (11), par l'introduction de nouvelles coordonnées, compare toutes les statistiques à la courbe de Gauss de formule

$$Y = 2e^{-\frac{x^2}{2}}$$

Cette courbe, indépendante de σ , peut être dessinée une fois pour toutes. Les nouvelles coordonnées sont les suivantes :

$$X = \frac{wx - b}{\sigma} \text{ et } Y = \frac{5}{wn} \sigma \cdot f(x)$$

Cette comparaison graphique de la concordance de deux courbes (progression arithmétique et progression géométrique sur l'axe des x) avec la courbe de Gauss a l'avantage d'être simple, immédiate, accessible à chacun.

2. RESULTATS

Nous examinerons successivement chacun des 11 collectifs étudiés :

a) *Indice de Tuxford des écoliers bernois*

Tuxford a établi empiriquement la formule suivante, permettant d'apprécier le développement des enfants :

$$\text{Indice} = \frac{\text{Poids. (381 - \text{âge})}}{54. \text{ taille}}$$

poids en kg., âge en mois et taille en cm.

La valeur de l'indice varie entre 78 et 168, tombant en dessous de 96 pour les enfants à développement insuffisant, dépassant 125 pour les enfants adipeux. *Lauener* (14) calcule depuis 10 ans cet indice sur les écoliers de la ville de Berne. Il a mis très obligeamment ses fiches à notre disposition. Pour éliminer les variations saisonnières et annuelles de l'indice, nous avons relevé sa valeur pendant la même période pour chaque enfant : soit les 5276 mesures effectuées en janvier et février 1946. Les tabelles suivantes résument nos calculs :

Index de Tuxford des écoliers bernois
Statistique arithmétique

Milieu des classes	Classes	f (x)	x	x. f(x)	x ² f(x)	Coordonnées de la courbe de Gauss	
						X	Y
102,5	78- 84	16	-3	- 48	144	-2,53	0,02
	85- 91	157	-2	- 314	628	-1,88	0,23
	92- 98	694	-1	- 694	694	-1,23	1,01
	99-105	1407	0	0	0	-0,58	2,05
	106-112	1388	1	1388	1388	0,07	2,02
	113-119	927	2	1854	3708	0,73	1,35
	120-126	432	3	1296	3888	1,37	0,63
	127-133	145	4	580	2320	2,03	0,21
	134-140	61	5	305	1525	2,68	0,09
	141-147	31	6	186	1116	3,33	0,04
	148-154	12	7	84	588	3,98	0,02
	155-161	4	8	32	256	4,64	0,006
	162-168	2	9	18	162	5,29	0,003

5276

= n

4687

= S x. f(x)

16417

= S x² f(x)

$$w = 7$$

$$b = \frac{7}{5276} \cdot 4687 = + 6,2$$

$$\sigma^2 = \frac{49}{5276} \cdot 16417 - 38 = 115$$

$$\sigma = \pm 10,7$$

$$\epsilon = \pm 0,14$$

$$M = 108,70$$

Statistique géométrique

Milieu des classes	Classes log x	f(x)	x	x · f(x)	x ² f(x)	Coordonnées de la courbe de Gauss	
						X	Y
2,03526	1,89209-1,91811	5	-5	- 25	125	-3,08	0,008
	1,91812-1,94414	42	-4	- 168	672	-2,46	0,06
	1,94415-1,97017	208	-3	- 624	1872	-1,84	0,32
	1,97018-1,99620	627	-2	-1254	2508	-1,22	0,95
	1,99621-2,02223	1206	-1	-1206	1206	-0,60	1,82
	2,02224-2,04826	1323	0	0	0	0,02	2,02
	2,04827-2,07429	1047	1	1047	1047	0,64	1,60
	2,07430-2,10032	511	2	1022	2044	1,26	0,77
	2,10033-2,12635	195	3	585	1755	1,88	0,30
	2,12636-2,15238	68	4	272	1088	2,50	0,10
	2,15239-2,17841	34	5	170	850	3,13	0,05
	2,17842-2,20444	8	6	48	288	3,75	0,01
	2,20445-2,23047	2	7	14	98	4,38	0,003

5276

= n

- 219

= S x · f(x)

13533

= S x² f(x)

$$w = 0,02603$$

$$b = \frac{0,02603}{5276} \cdot -219 = -0,00108$$

$$\log M_0 = 2,03526$$

$$\log M = 2,03526 - 0,00108 = 2,03418$$

$$\log \sigma^2 = \frac{0,00068}{5276} : 13533 - (0,00108)^2 = 0,00174$$

$$\log \sigma = 0,0417$$

$$\log e = 0,00057$$

$$\sigma = \begin{matrix} + 10,90 \\ - 9,91 \end{matrix}$$

$$e = \begin{matrix} + 0,14 \\ - 0,14 \end{matrix}$$

$$M = 108,19$$

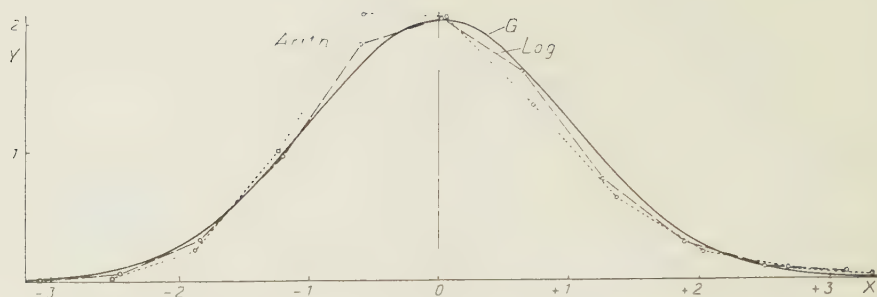


Fig. 1 : Index de Tuxford des écoliers bernois. G. : Courbe idéale de Gauss.

Arith. : Courbe de la statistique arithmétique.

Log. : Courbe de la statistique logarithmique ou statistique géométrique.

La courbe arithmétique est asymétrique, présentant un maximum fortement dévié vers les valeurs négatives de x . Ses branches s'éloignent sensiblement de la courbe idéale tant à droite qu'à gauche. La courbe géométrique a par contre une bonne symétrie et ses branches se rapprochent beaucoup mieux de la courbe de Gauss. L'avantage de la moyenne géométrique est évident dans ce cas. Il confirme ce que nous avons dit plus haut sur l'indication de la moyenne géométrique aux collectifs constitués par des rapports.

b) Sensibilité de la pulpe dentaire du lapin

Dolivo, de l'Institut de Physiologie de Lausanne, a mis à notre disposition 1608 mesures de seuil de sensibilité de la pulpe dentaire du lapin. Ce matériel n'a pas encore été publié. Les mesures sont effectuées selon la méthode de Rucksthul. Un trou de 2-3 mm. de profondeur est creusé sur la face latérale des deux incisives supérieures. Deux jours plus tard, on introduit des électrodes dans ces trous et on fait passer au travers des deux incisives des décharges de condensateurs. On mesure en volts le seuil douloureux qui se manifeste par des mouvements de mastication.

La statistique arithmétique donne les résultats suivants :

$$\sigma = + 7,5 \text{ volts}$$

$$\epsilon = \pm 0,19 \text{ volts}$$

$$M = 18,06 \text{ volts}$$

La statistique géométrique donne les résultats suivants :

$$\sigma = + 7,88 \text{ volts}$$

$$\sigma = - 5,37 \text{ volts}$$

$$\varepsilon = + 0,19 \text{ volts}$$

$$\varepsilon = - 0,18 \text{ volts}$$

$$M = 16,91 \text{ volts}$$

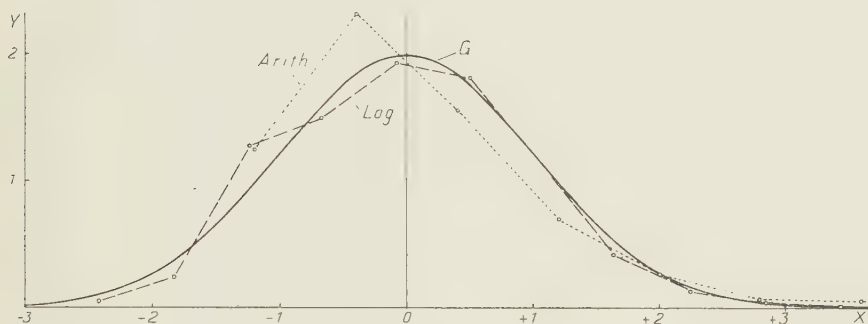


Fig. 2 : Sensibilité de la pulpe dentaire du lapin.

La meilleure concordance de la courbe logarithmique avec la courbe de Gauss idéale est évidente : la symétrie du maximum comme celle des valeurs extrêmes est fortement améliorée ; l'excédent du maximum arithmétique est réduit à la normale. Cette courbe confirme nos considérations sur l'emploi de la moyenne géométrique aux collectifs dépendant de la sensibilité du sujet d'expérience.

c) Sensibilité cutanée

Grandjean (13), à l'aide de l'esthésiomètre de Fleisch (9), a mesuré le seuil de sensibilité cutanée à la pression d'un même patient 5 fois par jour pendant plus de 4 mois dans des conditions expérimentales constantes. Selon l'échelle de l'appareil, les 686 mesures correspondent aux logarithmes de la pression exercée en mg. Elles nous donnent donc directement la statistique géométrique. Nous avons recherché les nombres de mg correspondant à ces mesures pour établir la statistique arithmétique qui donne les résultats suivants :

$$\sigma = \pm 248 \text{ mg}$$

$$\varepsilon = \pm 9,5 \text{ mg}$$

$$M = 846 \text{ mg}$$

Résultats de la statistique géométrique :

$$\sigma = + 288,0 \text{ mg}$$

$$\sigma = - 211,9 \text{ mg}$$

$$\varepsilon = + 9,2 \text{ mg}$$

$$\varepsilon = - 9,5 \text{ mg}$$

$$M = 803,5 \text{ mg}$$

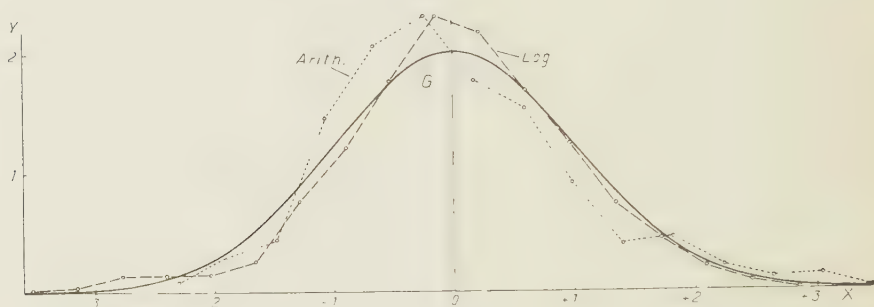


Fig. 3 : Sensibilité cutanée.

Bien que l'excédent du maximum persiste, la symétrie de la courbe est rétablie d'une manière très satisfaisante par la statistique géométrique. Ici aussi, l'avantage de la moyenne géométrique est évident et confirme son application aux collectifs dépendant de la sensibilité individuelle.

d) Dose mortelle de *Strophantus* pour la grenouille

Behrens (2) a publié les mesures de 149 doses mortelles de k-Strophantine pour la grenouille en injection intraveineuse d'une solution à 0.3 % à la vitesse de 0,01 cc. en 30 minutes. Les doses sont mesurées en γ par gr de grenouille.

La statistique arithmétique donne les résultats suivants :

$$\sigma = + 0,065 \gamma$$

$$\varepsilon = + 0,0053 \gamma$$

$$M = 0,326 \gamma$$

Résultats de la statistique géométrique :

$$\sigma = \begin{array}{l} + 0,072 \gamma \\ - 0,059 \gamma \end{array}$$

$$\varepsilon = \begin{array}{l} + 0,0054 \gamma \\ - 0,0053 \gamma \end{array}$$

$$M = 0,320 \gamma$$

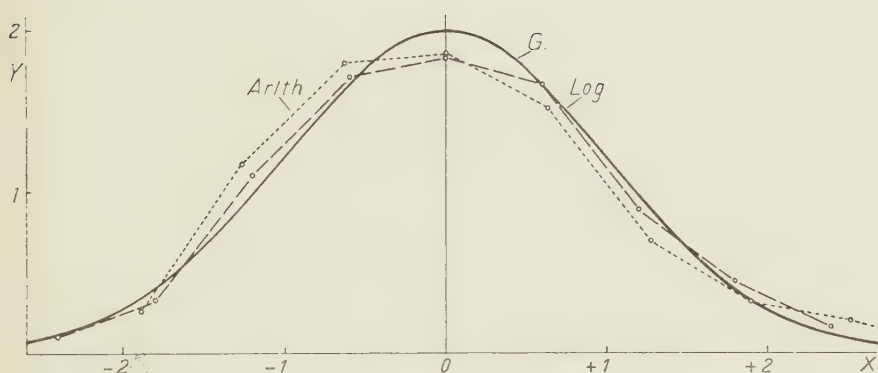


Fig. 4 : Dose mortelle de Strophantus pour la grenouille.

La meilleure symétrie de la courbe logarithmique justifie ici la moyenne géométrique et confirme son application, comme les deux exemples précédents, aux statistiques comprenant des mesures de sensibilité du sujet d'expérience.

e) Taux de cholestérol dans le sang

Georgi (12) a mis très obligeamment à notre disposition un matériel expérimental accumulé de 1930 - 1933 à la Clinique neuropsychiatrique de Breslau. Toutes les mesures que nous avons relevées ont été effectuées avec la même méthode, le matin et à jeun. Le matériel comprend 1310 mesures faites sur 439 malades.

La statistique arithmétique donne les résultats suivants :

$$\sigma = \pm 46,5 \text{ mg } \%$$

$$\varepsilon = \pm 1,28 \text{ mg } \%$$

$$M = 178,9 \text{ mg } \%$$

La statistique géométrique donne les résultats suivants :

$$\sigma = + 50,7 \text{ mg } \% \\ \quad - 39,1 \text{ mg } \%$$

$$\varepsilon = + 1,3 \text{ mg } \% \\ \quad - 1,1 \text{ mg } \%$$

$$M = 173,8 \text{ mg } \%$$

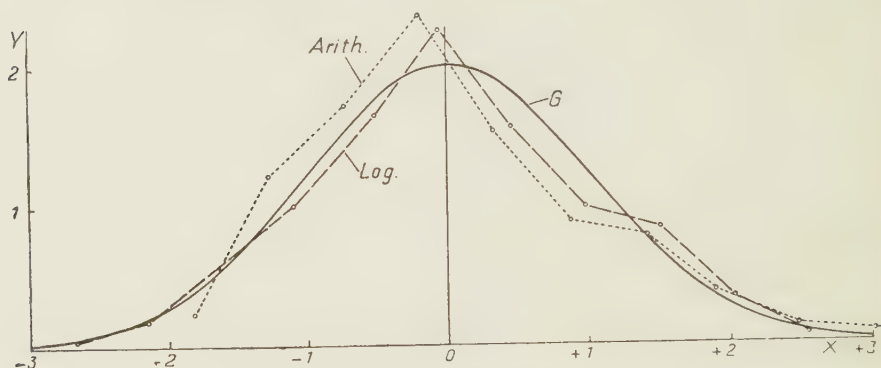


Fig. 5 : Taux de cholestérol dans le sang.

La meilleure symétrie du maximum, son abaissement, la meilleure concordance des branches de la courbe avec la courbe de Gauss, la meilleure symétrie des extrêmes parlent ici en faveur de la moyenne géométrique.

f) *Durée de la course de fond des conscrits*

Le Bureau fédéral de Statistiques (6) nous a obligeamment communiqué les temps mesurés pour la course de fond des conscrits de l'armée suisse lors du recrutement de 1944. Ce collectif comprend 29'454 mesures variant de 8'50" à 20'39". Le parcours doit être de 3000 m. Il faut cependant relever quelques différences entre les parcours des divers lieux de recrutement.

Voici les résultats de nos calculs pour la statistique arithmétique :

$$\sigma = \pm 1,03 \text{ min.}$$

$$\varepsilon_1 = \pm 0,006 \text{ min.}$$

$$M = \pm 12,58 \text{ min.}$$

et pour la statistique géométrique :

$$\sigma = + 1,05 \text{ min.}$$

$$\sigma = - 0,97 \text{ min.}$$

$$\varepsilon = + 0,006 \text{ min.}$$

$$\varepsilon = - 0,005 \text{ min.}$$

$$M = 12,52 \text{ min.}$$

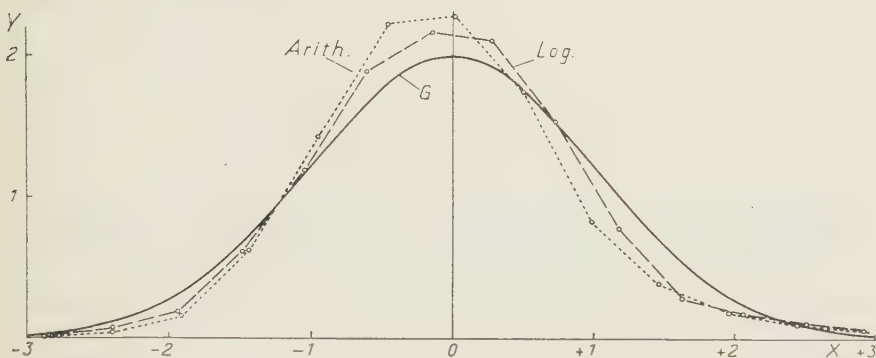


Fig. 6 : Durée de la course de fond des conscrits.

Tant l'amélioration de la symétrie de la courbe, que l'abaissement de son maximum, parlent ici aussi pour la moyenne géométrique. Il est intéressant de noter que c'est là la seule de nos statistiques comprenant des mesures de temps.

g) Taille des conscrits

Voici une statistique de 31'806 cas communiquée également par le Bureau fédéral de Statistique (6). Les mesures ont été faites au cm. près lors du recrutement de 1944. Statistique arithmétique :

$$\sigma = \pm 6,96 \text{ cm.}$$

$$\varepsilon = \pm 0,039 \text{ cm.}$$

$$M = 170,24 \text{ cm.}$$

Statistique géométrique :

$$\sigma = + 7,20 \text{ cm.}$$

$$\sigma = - 6,91 \text{ cm.}$$

$$\varepsilon = + 0,04 \text{ cm.}$$

$$\varepsilon = - 0,04 \text{ cm.}$$

$$M = 170,16 \text{ cm.}$$

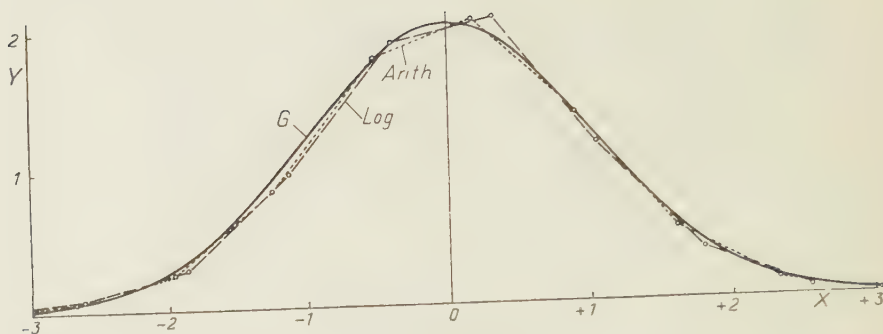


Fig. 7 : Taille des conscrits.

Comme la dispersion de ce collectif est faible, les courbes sont presque superposées. La courbe géométrique présente cependant une asymétrie vers les valeurs positives de x et un maximum excessif. Dans cet exemple, la moyenne arithmétique semble préférable.

h) Périmètre du thorax

Cette statistique comprend les 31'806 mesures du recrutement de 1944 communiquées par le Bureau fédéral de Statistique (6). Les mesures sont faites au cm. près. Résultats de la statistique arithmétique :

$$\begin{aligned}\sigma &= \pm 4,82 \text{ cm.} \\ \varepsilon &= \pm 0,03 \text{ cm.} \\ M &= 87,67 \text{ cm.}\end{aligned}$$

Statistique géométrique :

$$\begin{aligned}\sigma &= + 4,89 \text{ cm.} \\ &= - 4,59 \text{ cm.} \\ \varepsilon &= + 0,03 \text{ cm.} \\ &= - 0,02 \text{ cm.} \\ M &= 87,55 \text{ cm.}\end{aligned}$$

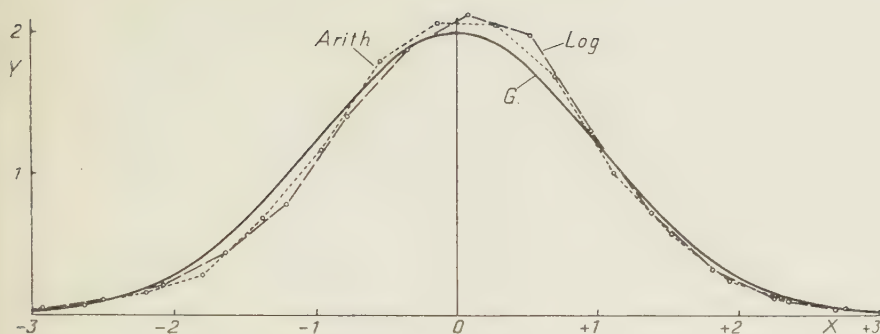


Fig. 8 : Périmètre du thorax des conscrits.

Comme dans la courbe précédente, la faible dispersion des mesures provoque une superposition presque complète des deux courbes. Ici aussi on remarque une légère asymétrie de la courbe logarithmique vers la droite ainsi qu'un maximum excessif. La statistique arithmétique semble préférable.

i) *Lancer en longueur*

On mesure la distance du jet d'une grenade de métal de 500 gr. Les 29'673 mesures du recrutement de 1944 ont été communiquées par le Bureau fédéral de Statistique (6). La précision des mesures se fait au mètre près.

Résultats de la statistique arithmétique :

$$\sigma = \pm 7,15 \text{ m.}$$

$$\varepsilon = \pm 0,04 \text{ m.}$$

$$M = 34,1 \text{ m.}$$

Résultats de la statistique géométrique :

$$\sigma = \begin{array}{l} + 8,59 \text{ m.} \\ - 6,82 \text{ m.} \end{array}$$

$$\varepsilon = \begin{array}{l} + 0,05 \text{ m.} \\ - 0,04 \text{ m.} \end{array}$$

$$M = 33,38 \text{ m.}$$

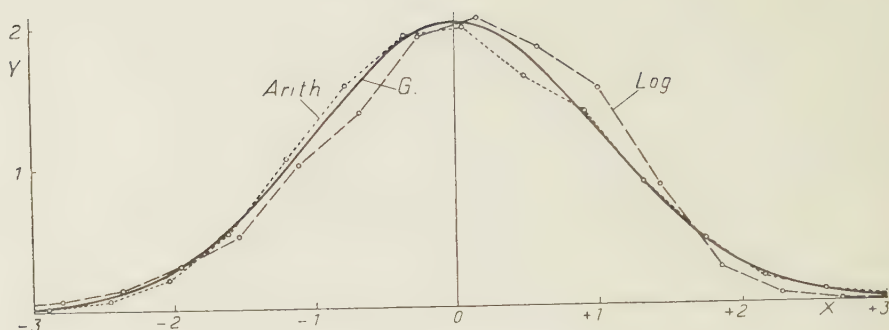


Fig. 9 : Lancer en longueur.

Bien que la courbe arithmétique présente ici une symétrie légèrement meilleure, ni l'une ni l'autre ne sont satisfaisantes. Il y aurait lieu de rechercher dans ce cas une moyenne intermédiaire entre la moyenne arithmétique et la moyenne géométrique comme l'ont fait *Fleisch et Tripod* (8) dans un collectif sur le temps de saignement du lapin.

k) Réfraction cristallinienne

Tron (20) a publié 200 mesures de réfraction cristallinienne chez l'homme.

La statistique arithmétique nous donne les résultats suivants :

$$\begin{aligned}\sigma &= \pm 2,75 \text{ dioptries} \\ \varepsilon &= \pm 0,19 \quad \gg \\ M &= 26,30 \quad \gg\end{aligned}$$

La statistique géométrique donne les résultats suivants :

$$\begin{aligned}\sigma &= \begin{array}{l} + 2,55 \\ - 2,84 \end{array} \text{ dioptries} \\ \varepsilon &= \begin{array}{l} + 0,20 \\ - 0,18 \end{array} \quad \gg \\ M &= 26,21 \quad \gg\end{aligned}$$

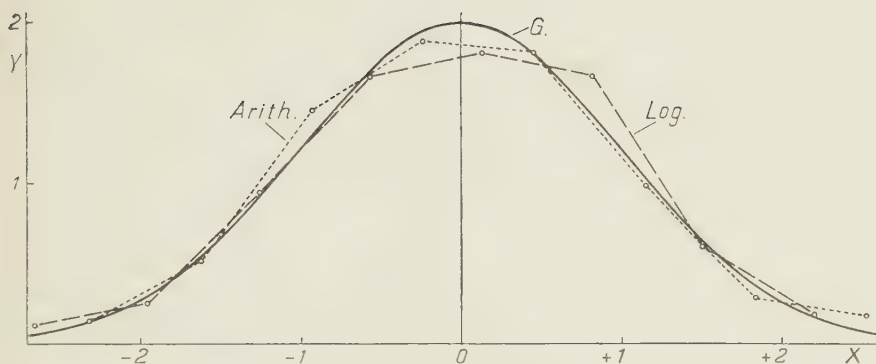


Fig. 10 : Réfraction cristallinienne.

Comme pour l'exemple précédent, la courbe idéale se trouve entre les courbes arithmétique et géométrique. Une moyenne intermédiaire entre la moyenne arithmétique et la moyenne géométrique est indiquée également dans ce cas.

1) Formule d'Arneth

Stephani (19) calcule la formule d'Arneth en comptant le nombre de noyaux de cent polynucléaires. Il a très obligeamment mis à notre disposition une statistique de 205 formules d'Arneth. Les valeurs oscillent entre 124 et 315. Les valeurs faibles correspondent à une « déviation vers la gauche », les grandes valeurs à une hypersegmentation.

Résultats de la statistique arithmétique :

$$\begin{aligned}\sigma &= \pm 36,5 \\ \varepsilon &= \pm 2,5 \\ M &= 205,2\end{aligned}$$

Résultats de la statistique géométrique :

$$\begin{aligned}\sigma &= + 33,7 \\ &\quad - 40,6 \\ \varepsilon &= + 2,7 \\ &\quad - 2,5 \\ M &= 201,8\end{aligned}$$

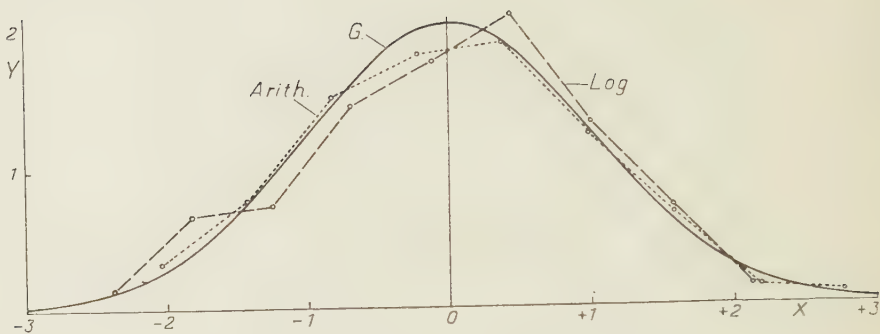


Fig. 11 : Formule d'Arneth.

L'asymétrie et le maximum excessif de la courbe géométrique semblent contre-indiquer cette statistique.

Résumé

Bien qu'en biologie on emploie presque exclusivement la moyenne arithmétique, celle-ci est souvent inadéquate. Dans certains cas la moyenne géométrique est plus juste. La courbe de Gauss établie sur la base du matériel statistique nous fournit un critère pour appliquer la moyenne arithmétique ou la moyenne géométrique à cette statistique. Il faut choisir pour l'axe des x une fonction telle que la symétrie de la courbe soit la meilleure possible. Si cette fonction est le logarithme des mesures, on applique ipso facto la moyenne géométrique.

Nous avons étudié 11 collectifs biologiques en comparant les courbes de leurs statistiques arithmétique et géométrique à la courbe de Gauss idéale. Six d'entre eux montrent une concordance plus satisfaisante pour la statistique géométrique. Ce sont :

1) un collectif composé de rapports, l'index de Tuxford. Nous avons déjà vu plus haut que pour des raisons de simple logique la moyenne géométrique s'impose dans ce cas.

2), 3), 4) Trois collectifs comprennent des mesures faisant intervenir la sensibilité du sujet d'expérience : sensibilité de la peau, sensibilité de la pulpe dentaire, dose mortelle d'un médicament. Ils confirment ainsi l'application à ces cas de la moyenne géométrique à laquelle nous avait incité l'étude de la dispersion de l'excitation et la loi de Weber-Fechner.

- 5) le taux sanguin du cholestérol.
- 6) un collectif de mesures de temps, la durée d'une course à pied semble aussi indiquer la moyenne géométrique.

Pour deux collectifs, la réfraction cristallinienne et la distance d'un lancer, une moyenne intermédiaire entre la moyenne arithmétique et la moyenne géométrique est indiquée comme dans l'exemple de *Fleisch et Tripod* sur le temps de saignement.

Enfin pour trois collectifs la moyenne arithmétique est préférable. Il s'agit chaque fois de nombres absolus : mesure de taille et de périmètre thoracique. nombre de noyaux de 100 leucocytes.

Bibliographie

- 1) *Arnold, E. H.* : J. med. Leysin, 22, 393 (1944).
- 2) *Behrens, B.* : Naunyn-Schmiedebergs Arch. 140, 240 (1929).
- 3) *Bristol-Roach, M.* : cité par Fisher (7), Ann. Bot. 39, 149 (1925).
- 4) *Charlier* : Vorlesung über Grundzüge der mathematischen Statistik, Hamburg (1920).
- 5) *Donzallaz, P.* : La statistique commerciale, Rouge et Co., Lausanne (1942).
- 6) *Bureau fédéral de Statistiques*. Annuaire statistique de la Suisse (1943) et communications spéciales.
- 7) *Fisher, R. A.* : statistical Methods for Research Workers, Oliver et Boyd, Londres (1936).
- 8) *Fleisch, A. et Tripod, J.* : Naunyn-Schmiedebergs Arch. 200, 135 (1942).
- 9) *Fleisch, A.* : Arch. f. Vererbungs-Forschung, Sozialanthropologie und Rassenhygiene, 18, 708 (1943) ;
Helv. Physiol. Acta 3, 351 (1945).
- 10) *Fleisch, A. et Posternak, J.* : Helv. Physiol. Acta 1, 421 (1943).
- 11) *Fueter, R.* : Das mathematische Werkzeug, Orell-Füssli, Zurich (1926).
- 12) *Georgi, F.* : Fortschr. d. Neurol. u. Psych. 1934, 340 ;
Schweiz. med. Wschr. 74, 539 (1944).
- 13) *Grandjean E.*, va paraître dans Helv. Physiol. Acta (1947).
- 14) *Lauener, P.* : Gesundheit u. Wohlfahrt, 12, (1944),
idem 2, (1947).
- 15) *Ludwig, H.* : Klin. Wschr., 20, 233 (1941).
- 16) *Melley, A.* : Une méthode capable de caractériser la variabilité en biologie, thèse Lausanne (1941).

- 17) *de Moncetz* : Initiation aux méthodes de la statistique, Libr. recueil Sirey 2me éd. (1941).
- 18) *de Montet*, Ch. et *Melley*, A. : Rev. med. Suisse rom., 60, 576, (1940).
- 19) *Stephani*, Th. : L'Avenir médical, 34, 150, (1937).
- 20) *Tron* : Graefes Arch. 122, 1.
- 21) *Yule*, G. U. et *Kendall*, M. G. : An Introduction to the Theory of Statistics, Griffin et Co, Londres (1940).

